Data File Caching Proxy Demonstrator

R.Brun, D. Duellmann, G. Ganis, A. Hanushevski, AJ. Peters

Demonstrator Goals

- prove expected benefits of proxy cache approach for analysis use case

  - better use of T2/T3 disk space and network bandwidth for popular data

    - transparent on-demand caching, cache hierarchies

  - low site management effort

    - cache only contains r/o data copies which can be refetched from origin SEs

Suggested Tests

- Modified xrootd server exists which maintains a local (disk/memory) cache

- Could be used with real analysis data models and access patters to evaluate

    - different cache implementations algorithms, eg

        - encode client read results via a unique hash for quick lookup

        - keep client vector reads in a sparse cache file + bitmap

    - different data replacement algorithms, eg

        - block based or file base cache ageing

        - user controlled cache expiry

- Ultimate implementation of proxy cache may be different

    - Eg SQUID, CDN

- but first measurements will help to constrain phase space to select suitable candidates.

Proposed Phases

- Phase 1 - July/August

  - Initial tests by IT-DSS and ROOT Team

  - measure cache efficiency and performance impact for LAN, 10ms, 100ms latency between origin and cache

    - need few sites volunteering to provide a test server (http or xroot)

    - experiment data files w different data models are available

- Phase 2 - September/October

  - experiment tests in a cache testbed

  - measure cache efficiency with realistic access pattern

- In contact with experiments